On the realism of tropical cyclone intensification in global storm-resolving climate models

Geophysical Research Letters

RESEARCH LETTER
10.1029/2024GL109841

On the Realism of Tropical Cyclone Intensification in Global Storm-Resolving Climate Models

Alexander J. Baker[1], **Benoît Vannière**[2], **and Pier Luigi Vidale**[1]

[1]National Centre for Atmospheric Science and Department of Meteorology, University of Reading, Reading, UK, [2]European Centre for Medium-Range Weather Forecasts, Bonn, Germany

Key Points:

- Simulated tropical cyclone characteristics analyzed in two fully coupled global climate models at atmospheric resolutions of 28 to 2.8 km
- Tropical cyclone intensification rate is close to observations at resolutions of 5 km or finer, and rapid intensification is captured
- Storm-resolving models also capture the observed relationship between high intensification rate and small inner-core size

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. J. Baker,
alexander.baker@reading.ac.uk

Citation:

Baker, A. J., Vannière, B., & Vidale, P. L. (2024). On the realism of tropical cyclone intensification in global storm-resolving climate models. *Geophysical Research Letters*, *51*, e2024GL109841. https://doi.org/10.1029/2024GL109841

Received 19 APR 2024
Accepted 6 AUG 2024

Author Contributions:

Conceptualization: Alexander J. Baker, Benoît Vannière, Pier Luigi Vidale
Formal analysis: Alexander J. Baker
Methodology: Alexander J. Baker, Benoît Vannière, Pier Luigi Vidale
Visualization: Alexander J. Baker
Writing – original draft: Alexander J. Baker
Writing – review & editing: Alexander J. Baker, Benoît Vannière, Pier Luigi Vidale

Abstract The physical processes governing a tropical cyclone's lifecycle are largely understood, but key processes occur at scales below those resolved by global climate models. Increased resolution may help simulate realistic tropical cyclone intensification. We examined fully coupled, global storm-resolving models run at resolutions in the range 28–2.8 km in the atmosphere and 28–5 km in the ocean. Simulated tropical cyclone activity, peak intensity, intensification rate, and horizontal wind structure are all more realistic at a resolution of ∼5 km compared with coarser resolutions. Rapid intensification, which is absent at typical climate model resolutions, is also captured, and exhibits sensitivity to how, and if, deep convection is parameterized. Additionally, the observed decrease in inner-core horizontal size with increasing intensification rate is captured at storm-resolving resolution. These findings highlight the importance of global storm-resolving models for quantifying risk and understanding the role of intense tropical cyclones in the climate system.

Plain Language Summary Simulating strong tropical storms (i.e., major hurricanes, super typhoons) with climate models is challenging because important processes that act to intensify a storm occur over spatial scales that are too small for global models to capture. Typical models lack sufficient resolution in the atmosphere and ocean, often constrained by computational resources. Recently, in a few models, resolution has increased to a point where each grid cell represents an area of just a few square kilometres, a significant leap of one or two orders of magnitude. We analyzed tropical storms simulated by these state-of-the-art, so-called storm-resolving models and found that peak tropical storm intensity and the rate at which storms intensify are both more realistic. These models also simulate the rapid intensification of tropical storms and capture the small eye diameters often seen in the most intense storms. Our work provides evidence that storm-resolving resolution may help us better understand the role of tropical storms in the climate system and predict their behavior in a warming climate.

1. Introduction

Tropical cyclones (TCs) are the costliest synoptic-scale meteorological phenomenon (World Meteorological Organization, 2021), containing mesoscale and convective-scale updrafts, eyewalls and rainbands. New analyses highlight an increasing trend in their destructiveness over recent decades (Elsner, 2020; Klotzbach et al., 2022) and the disproportionately high exposure of socioeconomically underprivileged populations to TC hazards globally (Jing et al., 2024). Simulating realistic TCs is therefore a vital ambition in weather and climate modeling. The physical processes governing a TC's lifecycle are relatively well understood, but many of these processes occur at scales below those resolved by global climate models. This undermines our understanding of TC risk and precludes reliable predictions of TC behavior in a warming climate. A key aspect of current climate model development for simulating TCs is increased horizontal resolution in both the atmosphere and ocean (Kreussler et al., 2021; Roberts et al., 2020; Vannière et al., 2020). State-of-the-art models, whose resolutions are approaching the kilometer scale, are termed global storm-resolving models (GSRMs; Satoh et al., 2019). It is anticipated that such resolutions—nominally <10 km—resolve many of the physical processes important for cyclogenesis and cyclone intensification (Judt et al., 2021).

Global climate models with horizontal resolutions of ∼1° fail to capture TC frequency (Roberts et al., 2020), intensity (Davis, 2018), and wind structure (Vannière et al., 2020). Increasing resolution from ∼1° to ∼0.25°—now typical in global modeling—improves simulated TC characteristics, but intensity remains significantly underestimated (Roberts et al., 2020). Furthermore, simulated intensification rates are too slow (Davis, 2018) and rapid intensification (RI) of TCs, defined as the 95th percentile of 24-hr intensity changes (Kaplan &

DeMaria, 2003), is not captured by coarse models. These shortcomings are also true of the latest, similar-resolution reanalyses (Aarons et al., 2021; Dulac et al., 2024). Capturing RI is particularly important because most of the strongest TCs undergo RI (Lee et al., 2016), yet forecast errors are typically several times larger for RI than non-RI cases (Majumdar et al., 2023; Trabing & Bell, 2020). Hurricane Otis (October 2023), for example, intensified by 74 m s^{-1}—from tropical storm to category 5—in under 24 hr prior to landfall, but this was not forecast by any operational model (García-Franco et al., 2024). In forecasting, initialization and assimilation of observational data aid model realism, but insufficient resolution contributes to intensity errors (Tam et al., 2021). GSRMs, run at resolutions akin to those of current operational forecasts, may help understand models' capabilities at these scales without initialization. In particular, model deficiencies may result from not capturing the observed association of the highest intensification rates with smaller TC inner-core size (Sparks & Toumi, 2022; Xu & Wang, 2015) and not representing inner-core processes (Trabing & Bell, 2020)—both require high atmospheric resolution (Moon et al., 2020). RI is expected to occur more frequently in a warmer climate (Bhatia et al., 2022), so the scientific and societal rationales for simulating realistic RI are equally compelling.

Judt et al. (2021) analyzed TCs in 40-day, atmosphere-only GSRM simulations and found that models capture realistic intensities, but exhibit disparate frequency, intensity, size, and structure biases. As noted by Judt et al. (2021), longer simulations are needed to draw general conclusions about simulated TC characteristics, and there is a need to assess storm-resolving resolution in the fully coupled system, as atmosphere-only models simulate more intense TCs because ocean feedbacks, which reduce intensity, are not included (Zarzycki, 2016). In this paper, we analyzed TCs in model-development simulations performed with two fully coupled GSRMs, spanning resolutions of 28 to 2.8 km in the atmosphere and 28 (0.25°) to 5 km (0.083°) in the ocean. These multi-year simulations provide an opportunity to examine TC realism at storm-resolving scales in detail ahead of forthcoming climate-length (i.e., ≥30 years) GSRM simulations. (Here, "realism" refers to similarity to observations.) Our focus is on simulated TC frequency, intensity, and horizontal wind structure. Model and observational data and methods are described in Section 2, results presented in Section 3, and further discussion presented in Section 4.

2. Data and Methods

2.1. Global Storm-Resolving Models

We analyzed simulations performed during model-development cycles 2 and three of the Next-Generation Earth Modeling Systems (nextGEMS) project (Table S1 in Supporting Information S1) with two fully coupled GSRMs. These are the European Center for Medium-Range Weather Forecast's (ECMWF) Integrated Forecasting System (IFS; Lang et al., 2023; Rackow et al., 2024), run with the Nucleus for European Modeling of the Ocean (NEMO) model at 0.25° resolution or the Finite Element Sea Ice Ocean Model (FESOM) on the eddy-resolving NG5 (5 km) grid, and the Max Planck Institute for Meteorology's Icosahedral Nonhydrostatic Weather and Climate Model (ICON; Hohenegger et al., 2023; Koldunov et al., 2023), run with the ICON-O ocean model (Korn et al., 2022). Changes between cycles 2 and 3 are as follows. In cycle 2, IFS cycle 47r3 was run at 9 km resolution (TC_O1279) with parameterized deep convection and NEMO, and at 4.4 and 2.8 km resolution (TC_O2559 and TC_O3999, respectively) with explicit deep convection and FESOM. In cycle 3, IFS cycle 48r1 was run with parameterized deep convection at 28 (TC_O399) and 9 km with NEMO and at 9 and 4.4 km with FESOM. In the 4.4-km simulation, cloud-base mass flux is scaled by a factor of 1/6, resulting in a more realistic distribution of intense precipitation (Rackow et al., 2024). In cycle 2, ICON was run with explicit convection at 10 and 5 km resolution, including a Smagorinsky turbulence parameterization. Another 10-km simulation was performed with the total turbulent energy (TTE) scheme of Axell (2002). In cycle 3, a Langmuir turbulence parameterization (Mauritsen et al., 2007) was included, which slightly deepens the ocean mixed layer. All simulations were initialized at 0 UTC on 20 January 2020, and simulation lengths are given in Table S1 of Supporting Information S1.

2.2. Tropical Cyclone Observations

Simulations were compared with version 4.0 of the National Oceanographic and Atmospheric Administration's International Best-Track Archive for Climate Stewardship (IBTrACS) data set (Knapp et al., 2018). All wind speed data were converted to 1-min "sustained" mean wind speed, following revised World Meteorological

Organization (WMO) guidance (Harper et al., 2010). The WMO-standard 10-min averaging period may be less comparable to instantaneous model output, and recent analyses (Judt et al., 2021; Roberts et al., 2020) also adopted a 1-min period. We prioritized wind data from the WMO Regional Specialized Meteorological Center responsible for each ocean basin. If no such data were available, we used data from the U.S. Department of Defense Joint Typhoon Warning Center or China Meteorological Administration, if available. In all cases, 0, 6, 12, and 18 UTC timesteps were used. Prior to the satellite era, IBTrACS data are less reliable (Schreck III et al., 2014; Vecchi et al., 2021) and operational data-collection and reporting procedures not only differ between WMO agencies but have also varied through time (Knapp & Kruk, 2010). Therefore, we used data over the period 1980–2022, and spur tracks (non-WMO tracks reported for a given TC, which may differ from the "main" track), short tracks with a lifetime of less than one day, and tracks lacking pressure or wind data were omitted from our analyses. This post-processing is similar to that performed by Dulac et al. (2024). Using post-1980 data minimizes the impact of discrepancies between operational delineation of observed tracks and the objective methods applied to model data. RI was defined using the observed threshold of 15.4 ms^{-1} $24hr^{-1}$ for all model simulations.

2.3. Tropical Cyclone Tracking

TCs were tracked in 6-hr nextGEMS data using the feature-tracking algorithm TempestExtremes (Ullrich & Zarzycki, 2017; Ullrich et al., 2021), chosen for its tendency to detect similar TC durations to IBTrACS (Bourdin et al., 2022). Following previous studies (Bourdin et al., 2022; Roberts et al., 2020; Ullrich et al., 2021), candidate points were identified by a closed-contour criterion: a sea-level pressure increase of ≥2 hPa over 5.5° from pressure minima. The presence of an upper-level warm core was identified by a geopotential thickness, Z_{250}–Z_{500}, decrease of ≥6 m (58.8 m^2 s^{-2}) over 6.5° from a reference Z_{250}–Z_{500} maximum within 1° of a given pressure minimum. Additional criteria were employed to merge candidate points, where necessary, within 6° and stitch candidates in time into cyclone tracks with a minimum duration of 54 hr and a minimum track length of 8°, which eliminates stationary features and spurious shallow lows. (All radii are geodesic.) Tracking algorithms typically diverge for weaker TCs (Bourdin et al., 2022), and, given recent work on the impact of horizontal resolution on TC intensity and structure (García-Franco et al., 2023; Judt et al., 2021; Kim et al., 2018; Roberts et al., 2020), tracking and attribution of sea-level pressure and near-surface (10-m) wind fields to tracked TCs here was performed at native model resolution to avoid smoothing out cyclone-associated features. For IFS, tracking was performed on dynamical fields remapped from the model's native (octahedral reduced Gaussian) grid to a regular grid, which does not affect tracking results (Figure S1 in Supporting Information S1). For ICON, TC tracking in cycle 2 was performed on ICON's native (icosahedral triangular) grid. In cycle 3, ICON model output in cycle three is on the equal-area, iso-latitude HEALPix hierarchical grid for which mesh size is given by 12×4^z, where z is a non-negative integer denoting the zoom level. Data were extracted at $z = 8$ and $z = 10$, corresponding to horizontal resolutions of 24 and 6 km, respectively, and remapped to a regular grid using nearest-neighbor interpolation. Example TempestExtremes commands are given in the Supporting Information S1. Multi-year simulations yield realistic global TC distributions, albeit with higher-than-observed TC frequency in the South Atlantic, but sample sizes in shorter simulations (1–2 years) result in noisy, less realistic distributions in some basins (Figure S2 in Supporting Information S1).

3. Results

3.1. Tropical Cyclone Frequency

Annual-mean TC frequency, n_{TC}, and TC days, d_{TC}, show sensitivity to resolution and differ between IFS and ICON (Figures 1a and 1b). In IFS, simulated n_{TC} in both cycle 2 and 3 is closer to observations at higher resolutions, although cycle-2 simulations at 4.4 and 2.8 km span less than 1 year. However, d_{TC} is underestimated by IFS, indicating the presence of shorter-lived TCs. Distributions of TC lifetime, t_{TC}, confirm this (Figure S3 in Supporting Information S1). Both models generally underestimate the frequency of t_{TC} in the range 10–17 days. IFS simulations at 4.4 and 2.8 km with explicit deep convection (cycle 2) differ from the other IFS and ICON simulations (Figure S3 in Supporting Information S1), and ICON simulates relatively short-lived TCs more frequently than observed, consistent with track density patterns (Figure S2 in Supporting Information S1). t_{TC} therefore exhibits greater sensitivity to whether deep convection is explicit or parameterized than to resolution. Additionally, cycle-3 IFS simulations at 9 km reveal an increase in n_{TC} and d_{TC} with FESOM at 5 km compared with NEMO at 0.25°. ICON shows stronger sensitivity to model formulation than to atmospheric resolution: simulated n_{TC} and d_{TC} exceed observational values in cycle 2 but are closer to IBTrACS in cycle 3. Accumulated

10.1029/2024GL109841

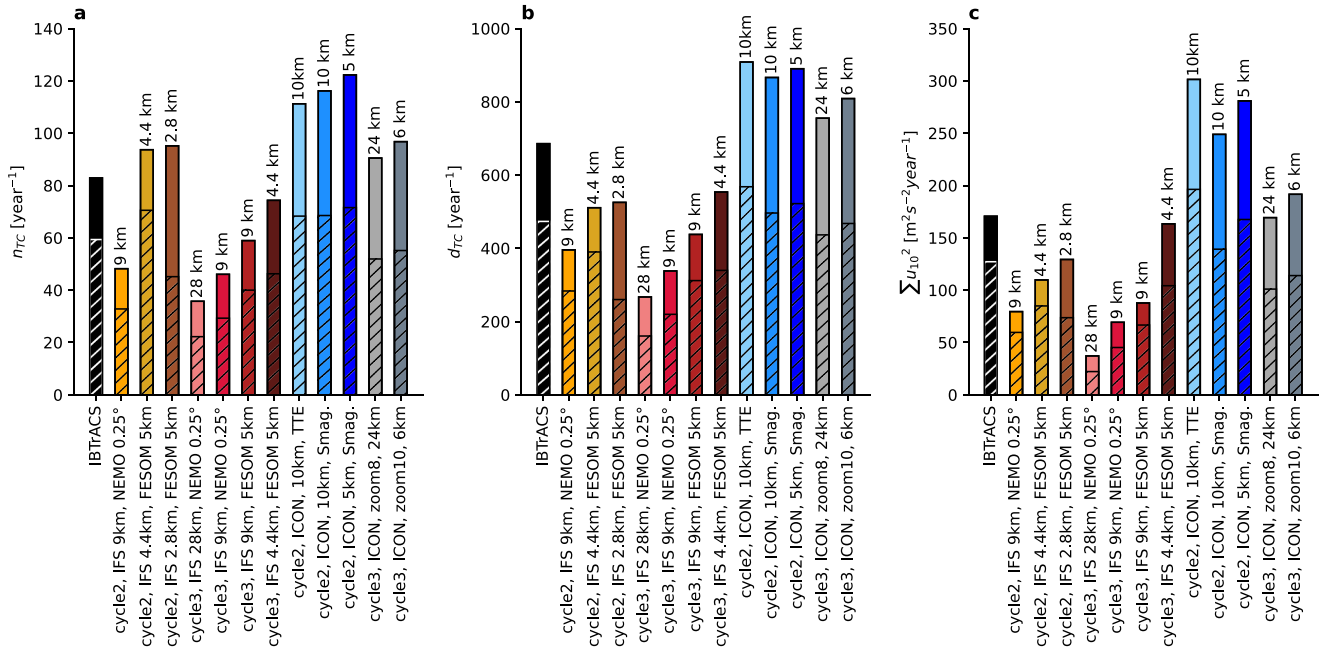

Figure 1. Annual-mean TC statistics in IBTrACS observations (black) and nextGEMS simulations: (a) TC frequency, n_{TC}, (b) cumulative TC days, d_{TC}, and (c) accumulated cyclone energy, $\sum u_{10}{}^2$, scaled by 10^{-4}. Bars represent global statistics and hatched areas represent the Northern Hemisphere. Tropical depressions (i.e., systems with $v_{max} < 17$ ms^{-1}) were excluded.

cyclone energy (ACE), a cumulative measure of TC activity (intensity and frequency), is well simulated in cycle 3 by IFS at 4.4 km and by ICON, although model performance is better globally than for the Northern Hemisphere (Figure 1c).

3.2. Tropical Cyclone Peak Intensity and Intensification Rate

A clear distinction between IFS and ICON exists in simulating the relationship between TC minimum central pressure, p_{min}, and near-surface wind speed, v_{max} (Figure 2). IFS generally captures the observed relationship and shows sensitivity to atmospheric resolution, particularly in cycle 3: for v_{max}, simulated TCs reach category 2 at 28 km and category 5 at 4.4 km. A caveat is the sample size of intense storms in a 5-year simulation may be insufficient to assess significance; the distinct curve of the 2.8-km simulation (cycle 2) is likely due to small sample size. The strongest v_{max} simulated by IFS is 89 m s^{-1}, which ranks among the strongest observed TCs. ICON simulates deep systems; the lowest simulated p_{min} is 863 hPa. However, v_{min} is too weak for a given p_{min}, and this model error deteriorates from cycle 2 to cycle 3. Marginal p_{min} distributions also demonstrate this. This is potentially related to the Langmuir turbulence parameterization used in cycle 3, which slightly deepens the ocean mixed layer, which may impact TC-induced sea-surface cooling (Wu et al., 2018). Two aspects of the observed wind-pressure relationship hamper comparisons with model data. First, observed wind speeds are estimates of maximum sustained gusts, whereas models output instantaneous wind. Second, the representativeness of point-wise observations, including retrievals and estimations, for simulated wind speeds, which are grid-cell averages, is questionable, especially at coarser resolution.

In addition to peak intensity, it is important to analyze intensification rates, including RI. We first composited the temporal evolution of v_{max} and p_{min} for the upper decile of observed and simulated TCs (Figures 3a and 3b). IFS exhibits strong sensitivity to resolution. At 28 km (cycle 3), TC intensity increases too slowly and peaks at approximately half the observed value. Simulations at 9 and 10 km are closer to IBTrACS. At 4.4 km, IFS mimics observations for composite v_{max} (Figure 3a) and p_{min} (Figure 3b). ICON at 5 km (cycle 2) performs well for v_{max} but p_{min} is too low, consistent with Figure 2. Unexpectedly, ocean model/resolution has little apparent impact in IFS: 9-km simulations with FESOM or NEMO show little difference for v_{max} or p_{min}.

RI is similarly sensitive to atmospheric resolution (Figures 3c and 3d). IFS at 2.8 km (cycle 2) and 4.4 km (cycle 3) simulates distributions of 24-hr intensification rates close to observations (Figure 3c). ICON at 5 km (cycle 2)

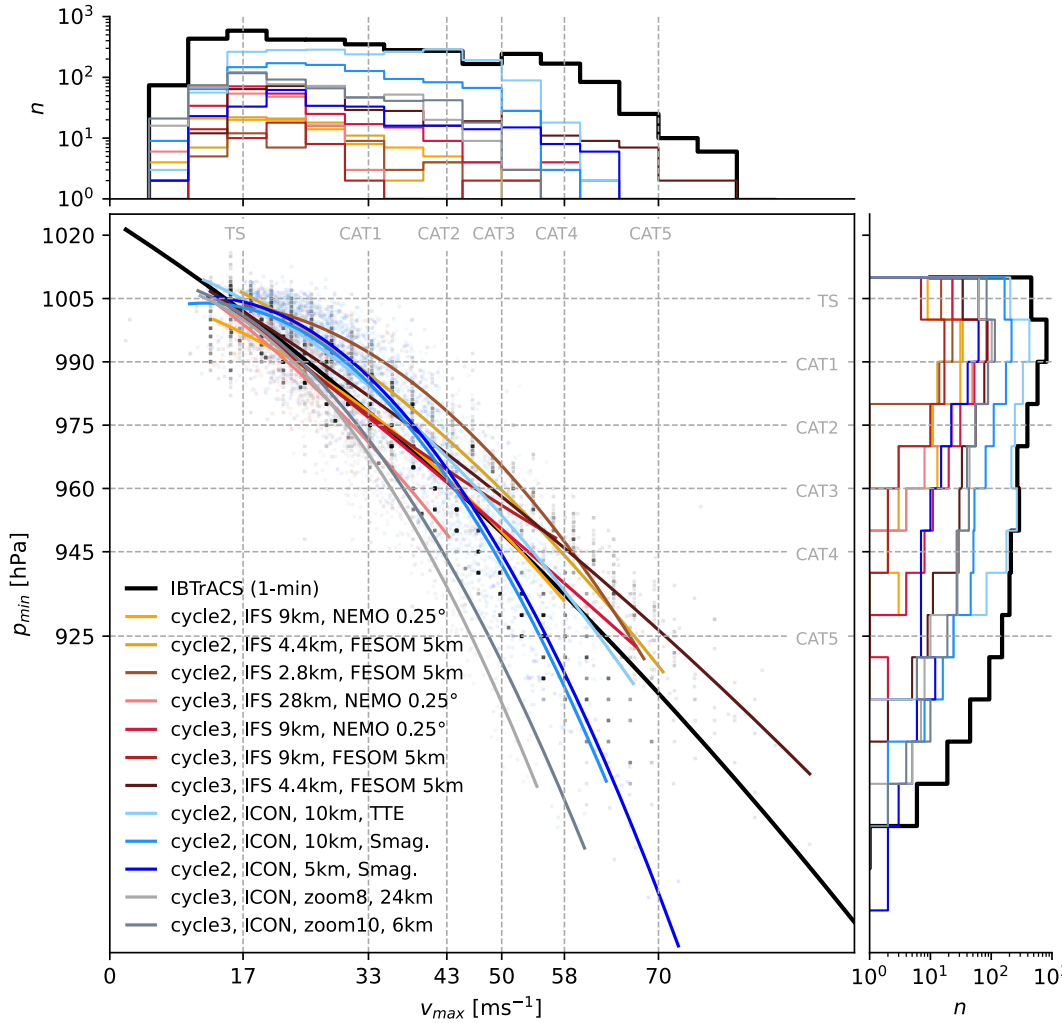

Figure 2. Relationship (second-order polynomial fits) between per-TC v_{max} and p_{min} for IBTrACS observations (black) and nextGEMS simulations, with marginal v_{max} and p_{min} frequency distributions shown in the top and right-hand panels, respectively (note log scales). TC categories, defined following Klotzbach et al. (2020) and the Saffir-Simpson scale for p_{min} and v_{max}, respectively, are delineated by gray, dashed lines. The legend gives the horizontal atmosphere and ocean resolutions of each simulation.

also simulates RI cases, though less frequently. Sample sizes are small at the upper tails of the distributions—hence, noisy model curves—and longer simulations will be particularly beneficial here. RI ratio, a measure of RI occurrence defined following Bhatia et al. (2022) as the ratio of 24-hr v_{max} changes above the RI threshold to non-RI v_{max} changes, simulated by IFS at 4.4 km (cycle 3) is close to the observed ratio (Figure 3d). In cycle 2, ICON simulates an RI ratio that is almost a factor 2 higher at 5 km than at 10 km, but remains below the observed value, and this metric deteriorates in cycle 3, again potentially due to a deeper mixed layer.

3.3. Tropical Cyclone Size and Wind Structure

Judt et al. (2021) reported variation among GSRMs in representing mean TC size. Here, we assess how realistic simulated inner-core, outer size, and horizontal wind structure are at high resolution, focusing on IFS simulations, which span the largest range in atmospheric resolution available in nextGEMS. Radial profiles of tangential wind speed, v_t, show smaller radius of maximum wind (RMW) with increasing resolution (Figure 4a). This result confirms the trend described by Vannière et al. (2020) and Moon et al. (2020), who demonstrated a reduction of horizontal inner-core size with resolution increases in the range ∼1 to ∼0.25°. Additionally, IFS simulations generally agree with the observed mean RMW across resolutions of 9–2.8 km. Again, little sensitivity to ocean

10.1029/2024GL109841

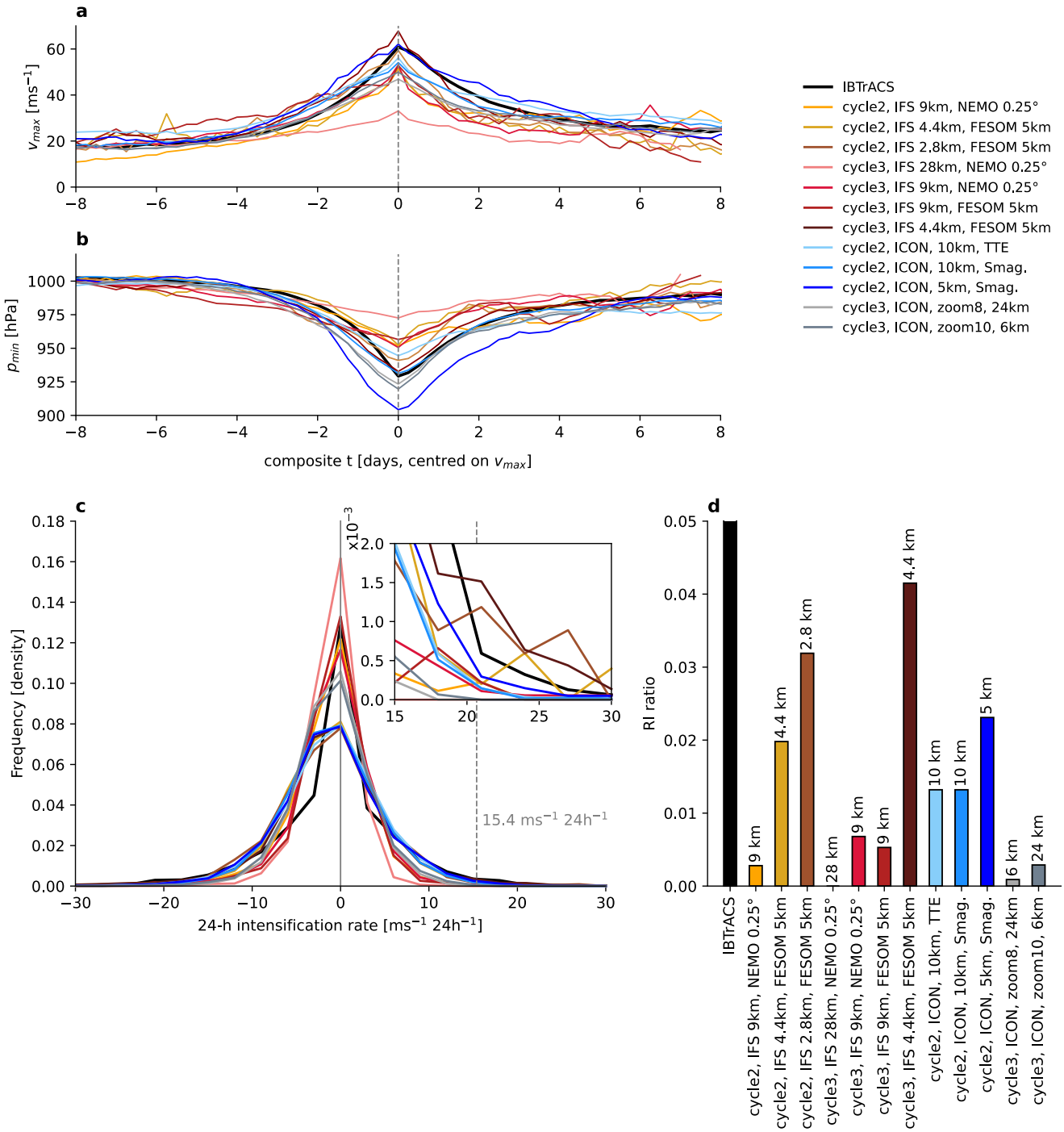

Figure 3. Composite TC intensity lifecycles for IBTrACS observations (black) and nextGEMS simulations, centered on peak v_{max} ($t = 0$), indicated by the gray, dashed line, for (a) v_{max} and (b) p_{min}. Analysis was performed for the upper decile (i.e., most intense by lifetime-maximum intensity) of both observed and simulated TCs. (c) Distribution of 24-hr v_{max} changes (positive = intensifying; negative = weakening), with (c, inset) intensification rates defined as RI (i.e., ≥15.4 m s^{-1} 24°hr^{-1}, indicated by the gray, dashed line) enlarged. For each TC, all RI instances were included. (d) Observed and simulated RI ratio, a measure of RI occurrence, computed for the 15.4 m s^{-1} 24 hr^{-1} threshold. By construction, the IBTrACS RI ratio is 0.05 (i.e., 5%). The legend gives the horizontal resolutions of the atmosphere and ocean for each simulation, and bar labels in (d) also give the atmosphere resolution.

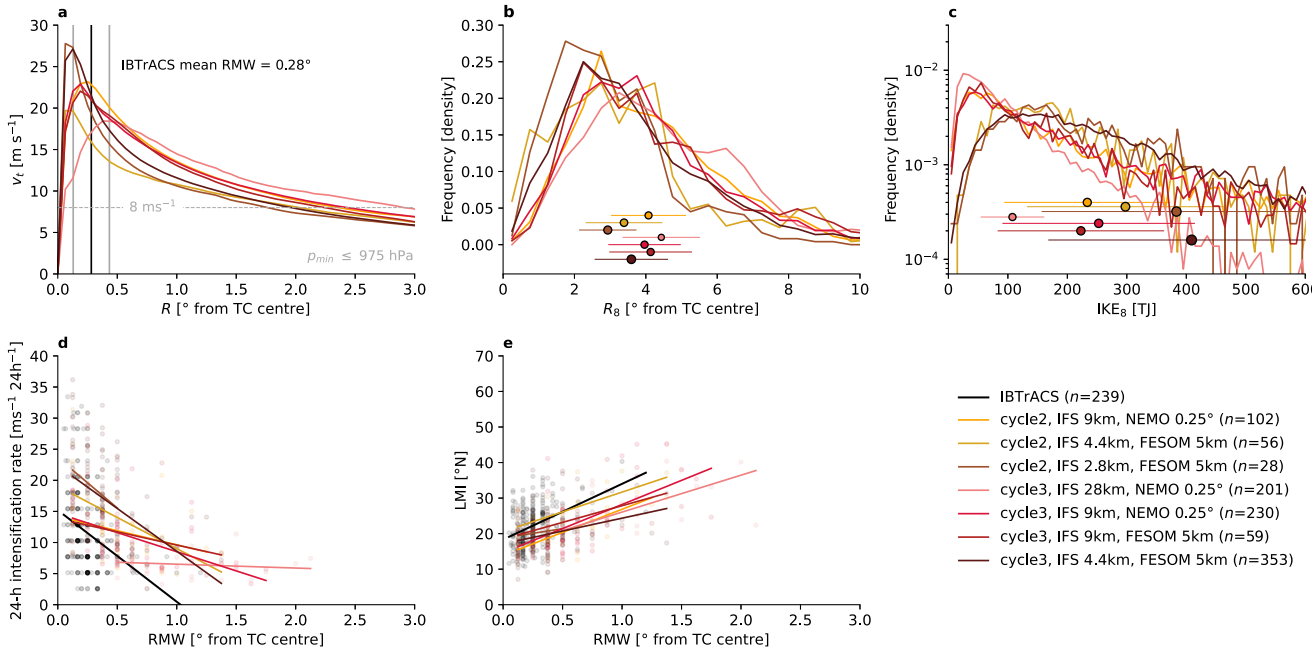

Figure 4. Horizontal wind structure of TCs simulated in IFS. (a) Azimuthal-mean (i.e., radial). profiles of v_t for all TCs with $p_{min} \leq 975$ hPa. The observed mean and standard deviation of radius of maximum wind (RMW) are indicated by the black and gray vertical lines, respectively. (b) Distribution of TC size, estimated by R_8. (c) Distribution of IKE, computed within R_8. Mean and standard deviation of R_8 and IKE$_8$ for each simulation are indicated by the jittered points (plotted against the x-axis only), with size scaled by mean v_{max}. (d) Relationship between TC intensification rate (24 hr^{-1}) and RMW, following Xu and Wang (2015). (e) Relationship between LMI and RMW. Linear fits are shown in (d) and (e). The legend gives the horizontal atmosphere and ocean resolutions of each simulation and the number of TCs, n, in each composite profile in (a); shorter simulations in cycle 2 have smaller n. For (b) and (c), all data were regridded to a common resolution of 0.25°.

model/resolution is seen between NEMO and FESOM runs with a 9-km atmosphere. Outer TC size, defined as the radius at which $v_t = 8$ m s^{-1}, R_8, decreases by a factor of ∼2 between resolutions of 28 and 2.8 km, from ∼3 to ∼1.7° (Figure 4a), and therefore smaller than observational estimates at 2.8 km (Schenkel et al., 2017, 2018). Frequency distributions of R_8 exhibit sensitivity to both resolution and the use of explicit convection: higher resolution reduces mean R_8, and explicit convection shifts the distribution to lower values (Figure 4b). Similarly, mean integrated kinetic energy within R_8, IKE$_8$, increases with increased resolution (although means are skewed by small high-IKE TC counts), and the distribution shifts toward higher values with either explicit convection or reduced cloud-base mass flux (Figure 4c), which increase inner-core v_t (Figure 4a). At lower resolution, simulated IKE$_8$ exceeds ∼100 TJ less frequently at 28 and 9 km than in simulations at 4.4 and 2.8 km. This result differs somewhat from an analysis of models spanning resolutions of ∼1 to ∼0.25° (Kreussler et al., 2021), which revealed IKE to be insensitive to this range in resolution because the larger horizontal size of TCs simulated at ∼1° compensates low wind speeds. Here, however, high-IKE TCs are poorly sampled—note IKE$_8$ is noisy above ∼400 TJ (Figure 4c)—and we are unable to assess significance.

Lastly, we investigate the relationship between intensification rate and size. IBTrACS data show an inverse relationship between intensification rate and RMW. IFS simulations at 4.4 and 2.8 km reproduce the observed slope of this relationship, with high intensification rate associated with small RMW (Figure 4d). 9-km simulations exhibit flatter slopes, and the 28-km simulation fails to capture the relationship, exhibiting a flat slope. The slope is more sensitive to resolution than to whether convection is explicit or parameterized: contrast, for example, cycle-2 simulations at 4.4 and 2.8 km. IFS simulations are offset from observations, indicating that, for a given RMW, the simulated intensification rate tends to be higher than observed. However, this may be partly due to linear fitting: linear fits are shown only to illustrate the scatter for IBTrACS and each simulation, and the relationship between intensification rate and RMW is evidently non-linear (Chan & Chan, 2018). IFS also simulates a similar relationship between the latitude of maximum intensity (LMI) and RMW to that seen in observations (Figure 4e); the latitudinal dependence of TC size is well captured but insensitive to resolution. Binning TCs into broad intensity categories indicates that both models simulate LMI over a similar range of latitudes to that observed (Figure S4 in Supporting Information S1). However, simulations with explicit convection tend to

simulate higher-latitude LMI, particularly in the Southern Hemisphere, which may impact the model RMW range.

4. Summary and Discussion

We analyzed TC activity (n_{TC}, d_{TC} and ACE), peak intensity, intensification rate, and horizontal wind structure in multi-year, fully coupled simulations performed with two GSRMs—IFS and ICON. Our analyses demonstrate that simulated TC characteristics improve significantly with an increase in resolution in the atmosphere from 28 km (IFS)/24 km (ICON)—equivalent to previous-generation, HighResMIP-type global models—to storm-resolving resolution of $\lesssim$5 km. Globally, realistic TC frequency and ACE are simulated in IFS (cycle 2 and 3) and ICON (cycle 3). Simulated activity for the Northern Hemisphere, however, is below that observed, particularly for ACE. At resolutions of $\lesssim$5 km, both models capture the observed range of TC peak intensities overall, but the simulated wind-pressure relationship is model- and resolution-sensitive. IFS exhibits sensitivity to the parameterization of deep convection and ICON simulates p_{min} that is too low for a given v_{max}. Further work will examine the sensitivity of the wind-pressure relationship to recent changes in model physics.

A key finding presented here is that GSRMs capture realistic TC intensification rates. Composite intensification in IFS at 4.4 km and ICON at 5 km mimic IBTrACS observations closely, and this is sensitive to resolution as well as models' treatment of deep convection. Particularly noteworthy is the ability of IFS at 4.4 km in capturing RI, which remains challenging in numerical weather prediction (Majumdar et al., 2023), and that it does so at a frequency close to that observed. This improvement is related in part to the use of reduced cloud-base mass flux, which increases RI ratio significantly versus explicit convection. In this simulation, deep convection may be considered partially parameterized, and this, we hypothesize, prevents convection from being triggered too abruptly in the presence of instability, which can occur in simulations with explicit convection. This may allow convective available potential energy to build before convection occurs. Testing this, and evaluating its importance versus that of resolution, is necessary future work.

A second important finding is that the simulated relationship between intensification rate and TC size at $\lesssim$5 km resolution resembles the observed relationship, indicating that GSRMs may help understand what role, if any, size has in determining intensification rate. Physical understanding of the size-intensification relationship is developing (Chan & Chan, 2018). As simulated horizontal inner-core structure is constrained by resolution, TC circulation becomes more realistic at finer resolution, with reduced RMW (Majumdar et al., 2023) and off-center vertical velocity maxima (Moon et al., 2020). Sparks and Toumi (2022) related a TC's p_{min} tendency to the ratio of its RMW and inflow or outflow speed, finding that this relationship is particularly important for higher observed intensification rates. This implies an upper limit on intensification rate for a given RMW and indicates that not capturing realistic RMW prevents models from simulating realistic intensification. Potentially, inner-core size is a physical constraint preventing coarse-resolution models from capturing observed intensification rates—and RI, observed cases of which frequently exhibit a "pinhole eye" (Olander & Velden, 2007). This is also true of reanalyses, which tend to overestimate inner-core size and underestimate outer-core size, although the representation of outer size is improved in ECMWF's Fifth Reanalysis, particularly in the eastern North Pacific basin where TCs are typically smaller than in other basins (Bian et al., 2021). Reanalysis TC intensity is more strongly correlated with observed TC size than with observed intensity or latitude (Schenkel & Hart, 2012), but reanalyses are unable to resolve the inner-core wind structures of intense TCs (Bian et al., 2021; Schenkel et al., 2017).

Shou et al. (2021) reported a decreasing trend in the inner-core size of Western North Pacific TCs since 1981, coinciding with an increasing intensity trend. However, coarse-resolution models and reanalyses are of limited use in explaining such trends. We found GSRMs capture RI more accurately but, crucially, also simulate a more realistic relationship between inner-core size and intensification rate. From a theoretical standpoint, the extent to which TC intensification rate depends on size remains an open question, but by reproducing the small RMW associated with the highest intensification rates and representing the observed range in intensity (and other characteristics), GSRMs may be an important tool in explicating observed TC trends.

Finally, we anticipate that other TC processes for which intensification rate plays a significant role may be simulated more realistically by GSRMs, particularly air–sea interactions, such as those between ocean eddies and developing TCs (Zhang et al., 2020), and TC-related precipitation extrema (Lamers et al., 2023). In our analysis, intensification rate is insensitive to ocean resolution, but previous case-study work, for example, found stronger cold wakes with increased ocean resolution (Polichtchouk et al., 2022). Further studies of TC interactions with the

ocean mixed layer at high-resolution will be valuable. Ongoing efforts to integrate GSRMs over multiple decades will allow these questions to be explored more fully, advancing our understanding of the role of intense TCs in the climate system and of their behavior in a warming climate.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Model data from the nextGEMS project are available via easy.gems.dkrz.de. IBTrACS data are available via ncei.noaa.gov/products/international-best-track-archive. *TempestExtremes* is available at climate.ucdavis.edu/tempestextremes and example commands are given in the Supporting Information S1. Data analysis and visualisation code is available from github.com/ncas-metoffice-hrcm/nextGEMS_TCs/.

Acknowledgments
AJB received support from the nextGEMS project (European Union Horizon, 2020 grant agreement 101003470). Paul Ullrich and Colin Zarzycki are thanked for supporting the authors' use of *TempestExtremes*. The authors are also grateful to two anonymous reviewers for their constructive feedback on the initial draft of this paper.

References

Aarons, Z. S., Camargo, S. J., Strong, J. D. O., & Murakami, H. (2021). Tropical cyclone characteristics in the MERRA-2 reanalysis and AMIP simulations. *Earth and Space Science*, *8*(3), e2020EA001415. https://doi.org/10.1029/2020ea001415

Axell, L. B. (2002). Wind-driven internal waves and Langmuir circulations in a numerical ocean model of the southern Baltic Sea. *Journal of Geophysical Research*, *107*(C11), 3204. https://doi.org/10.1029/2001jc000922

Bhatia, K., Baker, A., Yang, W., Vecchi, G., Knutson, T., Murakami, H., et al. (2022). A potential explanation for the global increase in tropical cyclone rapid intensification. *Nature Communications*, *13*(1), 6626. https://doi.org/10.1038/s41467-022-34321-6

Bian, G.-F., Nie, G.-Z., & Qiu, X. (2021). How well is outer tropical cyclone size represented in the ERA5 reanalysis data set? *Atmospheric Research*, *249*, 105339. https://doi.org/10.1016/j.atmosres.2020.105339

Bourdin, S., Fromang, S., Dulac, W., Cattiaux, J., & Chauvin, F. (2022). Intercomparison of four algorithms for detecting tropical cyclones using ERA5. *Geoscientific Model Development*, *15*(17), 6759–6786. https://doi.org/10.5194/gmd-15-6759-2022

Chan, K. T. F., & Chan, J. C. L. (2018). The outer-core wind structure of tropical cyclones. *Journal of the Meteorological Society of Japan*, *96*(4), 297–315. https://doi.org/10.2151/jmsj.2018-042

Davis, C. A. (2018). Resolving tropical cyclone intensity in models. *Geophysical Research Letters*, *45*(4), 2082–2087. https://doi.org/10.1002/2017gl076966

Dulac, W., Cattiaux, J., Chauvin, F., Bourdin, S., & Fromang, S. (2024). Assessing the representation of tropical cyclones in ERA5 with the CNRM tracker. *Climate Dynamics*, *62*(1), 223–238. https://doi.org/10.1007/s00382-023-06902-8

Elsner, J. B. (2020). Continued increases in the intensity of strong tropical cyclones. *Bulletin of the American Meteorological Society*, *101*(8), E1301–E1303. https://doi.org/10.1175/bams-d-19-0338.1

García-Franco, J. L., Gómez-Ramos, O., & Domínguez, C. (2024). Hurricane *Otis*: The costliest and strongest hurricane at landfall on record in Mexico. *Weather*, *79*(6), 182–184. https://doi.org/10.1002/wea.4555

García-Franco, J. L., Lee, C.-Y., Camargo, S. J., Tippett, M. K., Kim, D., Molod, A., & Lim, Y.-K. (2023). Climatology of tropical cyclone precipitation in the S2S models. *Weather and Forecasting*, *38*(9), 1759–1776. https://doi.org/10.1175/waf-d-23-0029.1

Harper, B. A., Kepert, J. D., & Ginger, J. D. (2010). *Guidelines for converting between various wind averaging periods in tropical cyclone conditions. WMO/TD-No. 1555*. World Meteorological Organization.

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., et al. (2023). ICON-sapphire: Simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, *16*(2), 779–811. https://doi.org/10.5194/gmd-16-779-2023

Jing, R., Heft-Neal, S., Chavas, D. R., Griswold, M., Wang, Z., Clark-Ginsberg, A., et al. (2024). Global population profile of tropical cyclone exposure from 2002 to 2019. *Nature*, *626*(7999), 549–554. https://doi.org/10.1038/s41586-023-06963-z

Judt, F., Klocke, D., Rios-Berrios, R., Vannière, B., Ziemen, F., Auger, L., et al. (2021). Tropical cyclones in global storm-resolving models. *Journal of the Meteorological Society of Japan. Ser. II*, *99*(3), 579–602. https://doi.org/10.2151/jmsj.2021-029

Kaplan, J., & DeMaria, M. (2003). Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Weather and Forecasting*, *18*(6), 1093–1108. https://doi.org/10.1175/1520-0434(2003)018<1093:lcorit>2.0.co;2

Kim, D., Moon, Y., Camargo, S. J., Wing, A. A., Sobel, A. H., Murakami, H., et al. (2018). Process-oriented diagnosis of tropical cyclones in high-resolution GCMs. *Journal of Climate*, *31*(5), 1685–1702. https://doi.org/10.1175/jcli-d-17-0269.1

Klotzbach, P. J., Bell, M. M., Bowen, S. G., Gibney, E. J., Knapp, K. R., & Schreck, C. J. III. (2020). Surface pressure a more skillful predictor of normalized hurricane damage than maximum sustained wind. *Bulletin of the American Meteorological Society*, *101*(6), E830–E846. https://doi.org/10.1175/bams-d-19-0062.1

Klotzbach, P. J., Wood, K. M., Schreck, C. J., III, Bowen, S. G., Patricola, C. M., & Bell, M. M. (2022). Trends in global tropical cyclone activity: 1990–2021. *Geophysical Research Letters*, *49*(6), e2021GL095774. https://doi.org/10.1029/2021gl095774

Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C., & Schreck, C. J. I. (2018). *International Best Track Archive for Climate Stewardship (IBTrACS) project, version 4, national centers for environmental information*. National Oceanographic and Atmospheric Administration. https://doi.org/10.25921/82ty-9e16

Knapp, K. R., & Kruk, M. C. (2010). Quantifying interagency differences in tropical cyclone best-track Wind speed estimates. *Monthly Weather Review*, *138*(4), 1459–1473. https://doi.org/10.1175/2009mwr3123.1

Koldunov, N., Kölling, T., Pedruzo-Bagazgoitia, X., Rackow, T., Redler, R., Sidorenko, D., et al. (2023). nextGEMS: Output of the model development cycle 3 simulations for ICON and IFS. *World Data Center for Climate (WDCC) at DKRZ*. https://doi.org/10.26050/WDCC/nextGEMS_cyc3

Korn, P., Brüggemann, N., Jungclaus, J. H., Lorenz, S. J., Gutjahr, O., Haak, H., et al. (2022). ICON-O: The ocean component of the ICON Earth system model—Global simulation characteristics and local telescoping capability. *Journal of Advances in Modeling Earth Systems*, *14*(10), e2021MS002952. https://doi.org/10.1029/2021ms002952

Kreussler, P., Caron, L.-P., Wild, S., Loosveldt Tomas, S., Chauvin, F., Moine, M.-P., et al. (2021). Tropical cyclone integrated kinetic energy in an ensemble of HighResMIP simulations. *Geophysical Research Letters*, *48*(5), e2020GL090963. https://doi.org/10.1029/2020gl090963

Lamers, A., Devi, S. S., Sharma, M., Berg, R., Gálvez, J. M., Yu, Z., et al. (2023). Forecasting tropical cyclone rainfall and flooding hazards and impacts. *Tropical Cyclone Research and Review*, *12*(2), 100–112. https://doi.org/10.1016/j.tcrr.2023.06.005

Lang, S., Schepers, D., & Rodwell, M. (2023). IFS upgrade brings many improvements and unifies medium-range resolutions. *ECMWF Newsletter*, *176*, 21–28.

Lee, C.-Y., Tippett, M. K., Sobel, A. H., & Camargo, S. J. (2016). Rapid intensification and the bimodal distribution of tropical cyclone intensity. *Nature Communications*, *7*(1), 10625. https://doi.org/10.1038/ncomms10625

Majumdar, S. J., Magnusson, L., Bechtold, P., Bidlot, J. R., & Doyle, J. D. (2023). Advanced tropical cyclone prediction using the experimental global ECMWF and operational regional COAMPS-TC systems. *Monthly Weather Review*, *151*(8), 2029–2048. https://doi.org/10.1175/mwr-d-22-0236.1

Mauritsen, T., Svensson, G., Zilitinkevich, S. S., Esau, I., Enger, L., & Grisogono, B. (2007). A total turbulent energy closure model for neutrally and Stably stratified atmospheric boundary layers. *Journal of the Atmospheric Sciences*, *64*(11), 4113–4126. https://doi.org/10.1175/2007jas2294.1

Moon, Y., Kim, D., Camargo, S. J., Wing, A. A., Sobel, A. H., Murakami, H., et al. (2020). Azimuthally averaged wind and thermodynamic structures of tropical cyclones in global climate models and their sensitivity to horizontal resolution. *Journal of Climate*, *33*(4), 1575–1595. https://doi.org/10.1175/jcli-d-19-0172.1

Olander, T. L., & Velden, C. S. (2007). The advanced Dvorak technique: Continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery. *Weather and Forecasting*, *22*(2), 287–298. https://doi.org/10.1175/waf975.1

Polichtchouk, I., Mogensen, K., Hadade, I., Hatfield, S., & Anantharaj, V. (2022). Greater atmospheric resolution improves prediction of tropical cyclones. *ECMWF Newsletter*, *172*, 8–9.

Rackow, T., Pedruzo-Bagazgoitia, X., Becker, T., Milinski, S., Sandu, I., Aguridan, R., et al. (2024). Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5/NEMOv3.4. *EGUsphere*, *2024*, 1–59.

Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vanniere, B., et al. (2020). Impact of model resolution on tropical cyclone simulation using the HighResMIP–PRIMAVERA multimodel ensemble. *Journal of Climate*, *33*(7), 2557–2583. https://doi.org/10.1175/jcli-d-19-0639.1

Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Düben, P. (2019). Global cloud-resolving models. *Current Climate Change Reports*, *5*(3), 172–184. https://doi.org/10.1007/s40641-019-00131-0

Schenkel, B. A., & Hart, R. E. (2012). An examination of tropical cyclone position, intensity, and intensity Life cycle within atmospheric reanalysis datasets. *Journal of Climate*, *25*(10), 3453–3475. https://doi.org/10.1175/2011jcli4208.1

Schenkel, B. A., Lin, N., Chavas, D., Oppenheimer, M., & Brammer, A. (2017). Evaluating outer tropical cyclone size in reanalysis datasets using QuikSCAT data. *Journal of Climate*, *30*(21), 8745–8762. https://doi.org/10.1175/jcli-d-17-0122.1

Schenkel, B. A., Lin, N., Chavas, D., Vecchi, G. A., Oppenheimer, M., & Brammer, A. (2018). Lifetime evolution of outer tropical cyclone size and structure as diagnosed from reanalysis and climate model data. *Journal of Climate*, *31*(19), 7985–8004. https://doi.org/10.1175/jcli-d-17-0630.1

Schreck, C. J., III, Knapp, K. R., & Kossin, J. P. (2014). The impact of best track discrepancies on global tropical cyclone Climatologies using IBTrACS. *Monthly Weather Review*, *142*(10), 3881–3899. https://doi.org/10.1175/mwr-d-14-00021.1

Shou, H., Li, T., Sun, Y., Wang, L., & Liu, J. (2021). Decreasing trend of Western North Pacific tropical cyclone inner-core size over the past decades. *Journal of Meteorological Research*, *35*(4), 635–645. https://doi.org/10.1007/s13351-021-0214-z

Sparks, N., & Toumi, R. (2022). The dependence of tropical cyclone pressure tendency on size. *Geophysical Research Letters*, *49*(19), e2022GL098926. https://doi.org/10.1029/2022gl098926

Tam, H.-F., Choy, C.-W., & Wong, W.-K. (2021). Development of objective forecast guidance on tropical cyclone rapid intensity change. *Meteorological Applications*, *28*(2), e1981. https://doi.org/10.1002/met.1981

Trabing, B. C., & Bell, M. M. (2020). Understanding error distributions of hurricane intensity forecasts during rapid intensity changes. *Weather and Forecasting*, *35*(6), 2219–2234. https://doi.org/10.1175/waf-d-19-0253.1

Ullrich, P. A., & Zarzycki, C. M. (2017). TempestExtremes: A framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geoscientific Model Development*, *10*(3), 1069–1090. https://doi.org/10.5194/gmd-10-1069-2017

Ullrich, P. A., Zarzycki, C. M., McClenny, E. E., Pinheiro, M. C., Stansfield, A. M., & Reed, K. A. (2021). TempestExtremes v2.1: A community framework for feature detection, tracking, and analysis in large datasets. *Geoscientific Model Development*, *14*(8), 5023–5048. https://doi.org/10.5194/gmd-14-5023-2021

Vannière, B., Roberts, M., Vidale, P. L., Hodges, K., Demory, M.-E., Caron, L.-P., et al. (2020). The moisture budget of tropical cyclones in HighResMIP models: Large-scale environmental balance and sensitivity to horizontal resolution. *Journal of Climate*, *33*(19), 8457–8474. https://doi.org/10.1175/jcli-d-19-0999.1

Vecchi, G. A., Landsea, C., Zhang, W., Villarini, G., & Knutson, T. (2021). Changes in Atlantic major hurricane frequency since the late-19th century. *Nature Communications*, *12*(1), 4054. https://doi.org/10.1038/s41467-021-24268-5

World Meteorological Organization. (2021). WMO atlas of mortality and economic losses from weather, climate and water extremes, (1970–2019). Retrieved from https://library.wmo.int/idurl/4/57564

Wu, L., Wang, R., & Feng, X. (2018). Dominant role of the Ocean mixed layer depth in the increased proportion of intense typhoons during 1980–2015. *Earth's Future*, *6*(11), 1518–1527. https://doi.org/10.1029/2018ef000973

Xu, J., & Wang, Y. (2015). A statistical analysis on the dependence of tropical cyclone intensification rate on the storm intensity and size in the North Atlantic. *Weather and Forecasting*, *30*(3), 692–701. https://doi.org/10.1175/waf-d-14-00141.1

Zarzycki, C. M. (2016). Tropical cyclone intensity errors Associated with lack of two-way ocean coupling in high-resolution global simulations. *Journal of Climate*, *29*(23), 8589–8610. https://doi.org/10.1175/jcli-d-16-0273.1

Zhang, Y., Zhang, Z., Chen, D., Qiu, B., & Wang, W. (2020). Strengthening of the Kuroshio current by intensifying tropical Cyclones. *Science*, *368*(6494), 988–993. https://doi.org/10.1126/science.aax5758